

Position Paper on Visualization  
Robert L. Galloway Jr.  
Professor of Biomedical Engineering  
Professor of Neurosurgery, Professor of Surgery  
Director, Center for Technology-Guided Therapy  
Vanderbilt University

It was fascinating to read the 1987 document. The boldness of “Workstations should be placed on the desks of each and every researcher to give them immediate access to local graphics capabilities. Every scientist and engineer should have a personal workstation, just as people who must drive for a living need access to cars. Workstations range in price from \$5,000 to \$100,000.” is breathtaking. I am also astonished at the statement: “The gigabit bandwidth of the eye/visual cortex...”, I know of no calculations that have been done which suggests that gigabit rates are the true bandwidth of the visual system.

In reviewing the recommendations from the 1987 workshop, it is with absolutely no malice. Rather, by comparing their projections to the present state, we should be both careful and humble about our ability to raise issues and point to what will be needed in the future.

Additionally, I looked up medical imaging issues from 1987. The American College of Radiology was stating that monitors were not acceptable for reading either chest films or mammograms. CT scans were 512x512 pixels and MRI scans were 256x256 pixels. Predictions were that CT scans would be 2048x2048 by 2000 and MR's would be 1024x1024. Image acquisition would be via electron beam computed tomography and would be able to acquire images fast enough to “stop” motion in the body.

In 1987 I made a name for myself at Vanderbilt by writing a computer program which allowed the display of images from a magnetic resonance imaging machine that did not require that you put the data tape back on the scanner. It displayed in a sterling 4 gray levels and took 2.5 hours to display the first image. Within a month I had that down to 7.5 minutes per frame and found that such an advance was publishable.

We now have at our fingertips the ability to put raster, vector and primitive images on a screen at blazing rates, we routinely update screens with raster graylevel images, surfaces rendered from triangulated data sets and overlaid “transparency” displays at 52 frames per second on megapixel displays. By using 32 bit display controllers we can use true 24 bit color while maintaining the speed of lookup grayscale in the alpha channel. Bus speeds and cache sizes allow us to optimize the display of grayscale by placing lookup tables in the graphics path. This allows us to achieve truer gray scale matching on 8 bit monitors than the very expensive >\$50k 10 bit grayscale monitors of the late 1980's and early 1990s.

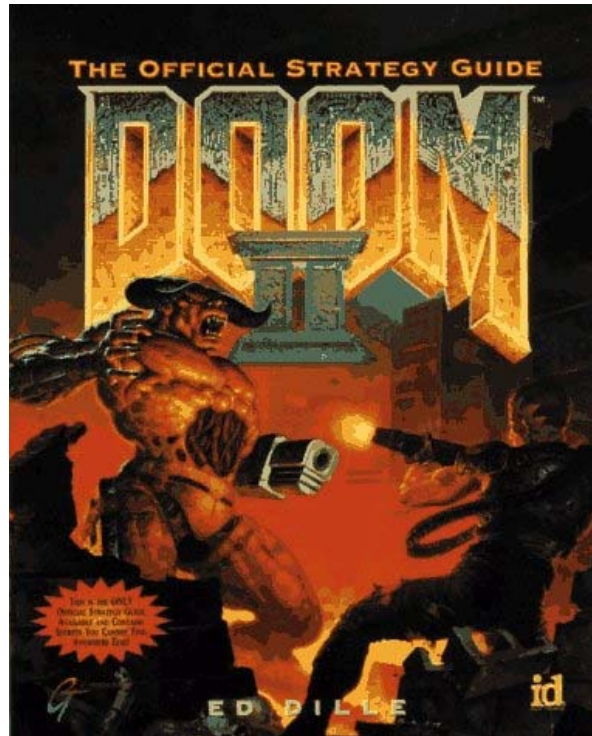
Graphic standards have evolved in a number of different ways but the ones that seem most important are the steady movement of primitive functions from software to firmware to hardware and the universal acceptance of at least those base standards.

Someone actively involved in the 1987 report but frozen in time would be astonished at the visualization computation firepower displayed on bright, high pixel count flat panels on our desks today. They would be flabbergasted at the \$2k-\$3k price tag.

They might be equally astonished to see that CT and MR images have remained relatively constant in their X,Y dimensions. CT's are still mostly 512x512 pixels and MR's are still 256x256. Where the changes have occurred include Z, that is interslice distance, 7 to 10 mm CT slices have become 0.7 to 1.2mm slice thicknesses and isotropically sampled MR's (where slice thickness is the same as pixel spacing) have become commonplace. Another place where images have changed is in time, multislice CT scanners and low flip angle MR sequences have reduced scan time without having to resort to exotic hardware such as EBCT or ultrahigh field MR with its concomitant potential for spatial distortions, not to mention bioeffects. Perhaps the biggest revolution in medical images has been the acceptance of computer monitors to read high contrast - range images such as chest films and mammograms. Vanderbilt is in the stages of removing the last film developers in the hospital and the light boxes to read films are not far behind. Part of that revolution is in the improvement of digital detection system so that images are not formed on film to begin with; but without acceptance that digital display methods were good enough, film would still be the method of choice. It has also been instructive to watch the changes in radiology resident training. Earlier, the residents were trained to "read the film" that is, all the data you are going to see is there. Now residents are taught to view the image dynamically; to use the display to change the contrast and range of the images to extract information from the data. PACS (picture archiving and communication systems) are making images available anywhere they are needed. If those locations mean the operating room and the radiology suite, a second copy can show instantly, simultaneously with no degradation in both places.

It is also of interest to note that this change has not gone unnoticed by industry. The leading PAC system companies include a roll-call of the major film companies: Kodak, Agfa and Fuji.

So now we take that 1987 expert and they might wonder what kind of things drove such a revolution. Was it their report? Did Congress see the light and provide unlimited funding for visualization development? Was there a visualization-driven breakthrough, such as cure for cancer, that caused all fields to embrace visualization as the way to go? There was one clear driving force but it (perhaps sadly) was not one of the above. It was:



Cover Photo: Doom II from Id Software.

The arrival of the “First Person Shooter” software changed the computer and visualization world in a fundamental way. FPS’s are heavily graphics oriented, demand little or no discernable display lag and drove sales of video hardware and computers. By demonstrating that a slight improvement in display performance lead to better game play, the FPS world heated up graphic hardware development and competition. ATI, Number Nine, Matrox, GeForce, nVidia and others all have ties to “scientific” computing but without a doubt the gamers have called the shots.

The field has also matured since 1987 by trying certain things and finding out what works and what doesn’t. This is especially true in the case of 3-D. Modern tomographic scanners have no problem acquiring 3-D data and the physician brains have no problem dealing with 3-D data but the visualization aspects of 3-D continue to elude us. Stereo techniques such as those with shutter glasses, autostereopsys, or heads up displays have made big splashes but have fallen into disuse. Very high tech (and high cost) immersion environments such as the CAVE system, have also failed to catch on in medicine.

The key to all of this seems to me to be a failure to understand the nature of vision, how visual information is presented and how it is processed. Vision is not 3-D, it might be best described as 2.5 D with stereo pair information giving us some idea as to where the perceived surface lies relative to our eyes. Vision can easily be overwhelmed especially if the viewer is merely the “rider” in the data stream instead of the “driver”. While heads-up displays have proven to be a boon to jet pilots, designers have had to scale back what they put on the screen to keep the pilot focused on flying.

Having said that, vision's capabilities and subtleties remain astonishing. Nothing is as diagnostic as hearing a surgeon say in an operating room, "that just doesn't look right." A biopsy sample sent to pathology from that site is rarely found to be normal. A chest image viewed by an experienced radiologist is rarely visible for 5 seconds before they have found what is wrong. The dictation and the gathering of additional data to confirm a diagnosis often take 20 times as long as that first viewing. In fact, studies of mammographer's eyes have shown that false positive rates go up (while true positive rates stay the same) when their dwell time (time spent looking at any given part of the image) goes up. When the forebrain overrules the visual cortex and associative areas, errors increase.

I design systems for image-guided therapy: surgery, ablation, brachytherapy, direct injection gene therapy, direct injection chemotherapy and the guided placement of neuroprosthetics and neuroprotective drugs. Such image-guided procedures can be considered eight-dimensional. That is, the three-space location (X,Y,Z) of some site of surgical interest (called a target), the trajectory orientation (Yaw, Pitch and Roll), Modality and Time. Modality refers to the distinct, and often orthogonal (in a contrast sense rather than a spatial sense) information possessed in various imaging modalities: CT, MRI, PET, SPECT, ultrasound, video images, fluoroscopes and angiograms. A user of a multimodal IGS system must traverse the data sets and resolve the underlying physiological state or function that gives rise to the imaging data presented on the screen.

I face two types of challenges in this process. The first is almost philosophical and yet is at the heart of visualization. – We have the ability to display virtually anything, what we have to figure out is what we *should* display. No computer visualization system in the world will be able to handle an 8 dimensional problem, and even if it did, our eyes and visual cortex couldn't process such information. However, the task of surgery can be reduced in dimension: If I show the surgeon a preplanned path and only ask him or her to stay on the path, then I "straighten out" a 3-D world to 1-D one. Airplanes attempting to land don't have to deal with Mode but the other 7 dimensions are real. Right time, right airport, right runway, upside down is still a bad outcome. So the idea of presenting the right information for the task at the right time in an intuitive fashion is a challenge. I intend to work with visual psychologists on this and would like to see additional support for cross-disciplinary teams.

The second challenge is much more mundane but well within grasp of government organizations. We bring a huge amount of visual data to the surgical task: CT, MRI, PET, SPECT, ultrasound, video images, fluoroscopes and angiograms. The development of the DICOM standard has made the importation of such information much easier. However, there still exist way too many "flavors" of DICOM to make the process universal to all medical settings.

Secondly, as surgical procedures move to less invasive techniques, intraoperative imaging is becoming more important. If one wishes to be able to use any manufacturer's video output into a guidance system you are stuck with S-video (if you are lucky) and NTSC (if you are not). That means jitter, field tear and anisotropic scaling are constant

issues. The development of a rational, national digital video standard for medical devices would greatly enhance the process. Such a standard would allow easy integration of intraoperative images into processed data streams and would allow any new video source to be integrated almost immediately. It would also allow standardized calibration processes so that anything seen could be localized.

If such a standard was adopted then the chip sets to accumulate and transfer the data could be put into production and added to any new device allowing for a gradual migration to the new standard. Innovative companies wishing to add value to their product would lead the way forcing the field to follow.